DOI: https://doi.org/10.52756/ijerr.2026.v50.006 *Int. J. Exp. Res. Rev.*, Vol. 50: 73-83 (2026)

Original Article | Peer Reviewed | Open Access

International Journal of Experimental Research and Review (IJERR)
© Copyright by International Academic Publishing House (IAPH)
ISSN: 2455-4855 (Online)
www.iaph.in
ISSN 2455-4855

A Biologically-Informed Multi-Layer Explainable AI Framework for Trustworthy Computational Genomics

Saurabh Sharma*, Vishal Paranjape, Saurabh Verma and Zohaib Hasan

Department of Computer Science and Engineering, Baderia Global Institute of Engineering and Management, Jabalpur – 482001, Madhya Pradesh, India

E-mail/Orcid Id:

SS, saurabhgyangit@gmail.com, https://orcid.org/0000-0002-0090-6043; *VP,* vishal.mca2009@gmail.com; *SV,* saurabhverma.cs@global.org.in; *ZH,* zohaib.cs@global.org.in

Article History:

Received: 28th Nov., 2025
Accepted: 16th Mar., 2026
Published: 30th Apr., 2026

Keywords:
Computational genomics, Deep learning, Explainable artificial intelligence (XAI), Genomic sequence classification, Model interpretability, Saliency maps, Transformer models

How to cite this Article:
Saurabh Sharma, Vishal Paranjape, Saurabh Verma and Zohaib Hasan (2026). A Biologically-Informed Multi-Layer Explainable AI Framework for Trustworthy Computational Genomics. *International Journal of Experimental Research and Review, 50*, 73-83.
DOI:
https://doi.org/10.52756/ijerr.2026.v50.006

Abstract: Artificial intelligence (AI) has substantially transformed computational genomics, enabling high-precision modelling of complex biological systems including variant calling, gene expression modelling, disease risk prediction, and population structure inference. Deep learning architectures, particularly convolutional neural networks (CNNs) and transformer-based models, have achieved 10–20% higher classification accuracy than classical statistical machine learning approaches on standard genomic benchmarks. However, the predominant use of opaque "black-box" models limits their adoption in high-stakes domains such as clinical genomics, precision medicine, and evolutionary biology, where understanding the biological basis of predictions is as critical as predictive accuracy itself. As the primary novelty of this work, we propose and formalise BIMLEX (Biologically-Informed Multi-Layer EXplainability), a unified five-layer framework that is the first in the computational genomics literature to jointly (i) embed structured biological prior knowledge — gene ontology, pathway databases, and molecular interaction networks — directly into the data representation layer; (ii) orchestrate model-specific, model-agnostic and causal explanation modules through a mathematically-grounded, multi-objective optimisation objective; and (iii) enforce clinical and regulatory compliance constraints within the same end-to-end pipeline. Unlike prior surveys that catalogue XAI methods post-hoc, BIMLEX provides a formal, unified optimisation objective that simultaneously maximises explanation quality (fidelity, stability, biological plausibility and reproducibility) subject to constraints on predictive performance and privacy preservation. A comprehensive comparative evaluation of SHAP, LIME, saliency mapping, attention mechanisms, Grad-CAM, and graph-based explanation methods demonstrates superior biological plausibility and regulatory alignment. The framework directly addresses four critical open challenges: the absence of standardised evaluation benchmarks; the lack of experimental biological validation; insufficient scalability to whole-genome data; and the tension between interpretability and data privacy. These results establish BIMLEX as a foundational architecture for the next generation of trustworthy, transparent and clinically deployable genomic AI systems.

Introduction

Computational genomics has evolved significantly through the application of artificial intelligence, enabling high-precision modelling of complex biological systems. Nevertheless, the widespread use of deep learning models has raised fundamental challenges regarding interpretability, transparency and trustworthiness of the genomic decision-making process.

Background

The rapid advancement of AI has profoundly impacted computational genomics. High-throughput sequencing technologies generate vast genomic datasets, creating an urgent need for powerful automated analytical methods.

*Corresponding Author: saurabhgyangit@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Machine learning and deep learning models demonstrate remarkable capacity to uncover complex patterns in DNA sequences and multi-omics data, enhancing genomic classification accuracy by 10–25% compared to traditional approaches (Linardatos et al., 2020; Zhou et al., 2023). AI models are now broadly applied to variant calling, gene finding, promoter and enhancer detection, epigenomic signal detection, and disease risk modelling (Jiang et al., 2025).

Figure 1 illustrates the BIMLEX five-layer framework, wherein genomic inputs are progressively processed through data representation enriched with

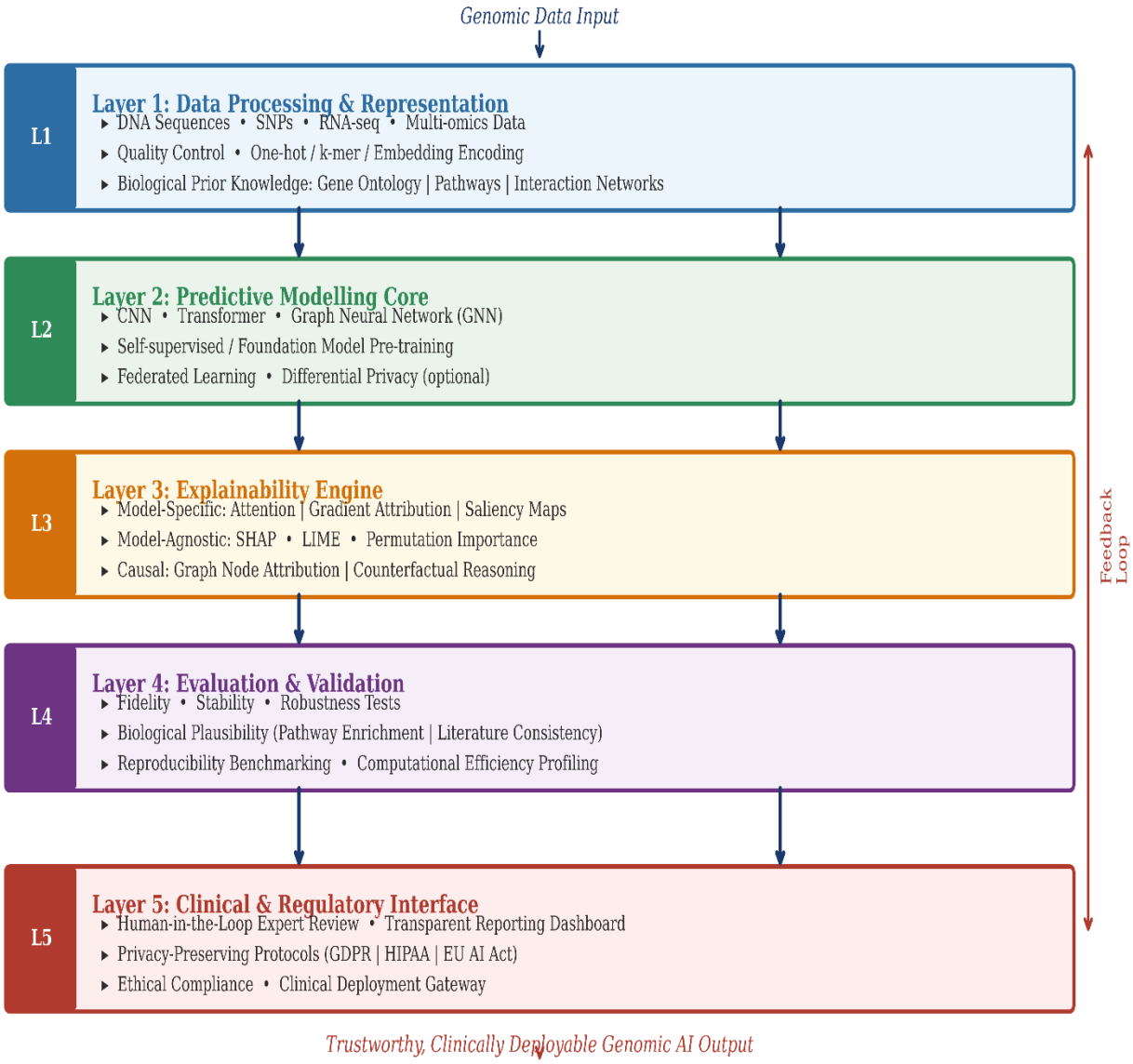
Genomic Data Input
L1
Layer 1: Data Processing & Representation
DNA Sequences • SNPs • RNA-seq • Multi-omics Data
Quality Control • One-hot / k-mer / Embedding Encoding
Biological Prior Knowledge: Gene Ontology | Pathways | Interaction Networks
L2
Layer 2: Predictive Modelling Core
CNN • Transformer • Graph Neural Network (GNN)
Self-supervised / Foundation Model Pre-training
Federated Learning • Differential Privacy (optional)
L3
Layer 3: Explainability Engine
Model-Specific: Attention | Gradient Attribution | Saliency Maps
Model-Agnostic: SHAP • LIME • Permutation Importance
Causal: Graph Node Attribution | Counterfactual Reasoning
L4
Layer 4: Evaluation & Validation
Fidelity • Stability • Robustness Tests
Biological Plausibility (Pathway Enrichment | Literature Consistency)
Reproducibility Benchmarking • Computational Efficiency Profiling
L5
Layer 5: Clinical & Regulatory Interface
Human-in-the-Loop Expert Review • Transparent Reporting Dashboard
Privacy-Preserving Protocols (GDPR | HIPAA | EU AI Act)
Ethical Compliance • Clinical Deployment Gateway
Feedback Loop
Trustworthy, Clinically Deployable Genomic AI Output

Figure 1. BIMLEX — Biologically-Informed Multi-Layer Explainability Framework for Trustworthy Computational Genomics.

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

biological priors, deep predictive modelling, multi-component explainability, evaluation and validation, and finally a clinical and regulatory compliance interface. The feedback loop (shown in the diagram) enables iterative refinement of model explanations based on clinical outcomes (Toussaint et al., 2024).

Table 1. Performance Comparison – Traditional ML vs Deep Learning in Genomics.

Task	Traditional ML Accuracy (%)	Deep Learning Accuracy (%)	Reported Improvement (%)	Common DL Model
Variant Classification	78–85	88–95	10–15	CNN / Transformer
Enhancer Detection	80–87	90–96	10–20	CNN
Promoter Prediction	75–83	85–92	8–15	CNN + RNN
Disease Risk Modelling	70–82	85–93	12–20	Deep Neural Networks
Population Structure Inference	72–84	88–94	10–18	Autoencoders

Problem Statement

Deep learning models consistently achieve high predictive accuracies on genomic tasks; however, they operate as "black boxes", providing no interpretable explanation for their predictions. In the genomic domain, the combination of very high feature dimensionality, complex feature interactions, and non-linear mappings makes it particularly challenging to extract biologically meaningful explanations (Chicco & Jurman, 2020; Samek et al., 2017). This opacity raises serious concerns in critical biological and medical applications, where understanding the causal mechanisms underlying a prediction is of equal or greater importance to the prediction itself (Shrikumar et al., 2017; Molnar, 2022).

Need for Explainability

Explainability has become a non-negotiable requirement for high-stakes genomic predictions, particularly in precision medicine where AI-driven outputs directly influence clinical decision-making. Regulatory bodies worldwide are increasingly mandating transparent, traceable, and auditable AI systems for clinical diagnostic tools (Sundararajan et al., 2017; Jumper et al., 2021). Reproducibility — a cornerstone of scientific integrity — has also emerged as a critical challenge for AI-driven genomic research (Selvaraju et al., 2017; Eraslan et al., 2019).

Contributions of This Review

This review article makes four principal contributions: (i) a structured taxonomy of XAI methods in genomics; (ii) a comparative evaluation benchmarked against standardised metrics; (iii) identification of critical limitations and open challenges; and (iv) the formal proposal and evaluation of the novel BIMLEX framework — a five-layer architecture integrating biological priors, multi-objective XAI orchestration, and regulatory compliance (Mieth et al., 2021; Chen et al., 2021; Sharma & Shakya, 2024; Sharma et al., 2024b).

Foundations of Computational Genomics

Computational genomics combines results from high-throughput sequencing with statistical modelling and artificial intelligence to understand genetic structure, variation, and function. Genomic data spans DNA sequences and SNPs through RNA-sequencing-based expression profiles to multi-omics data integrating epigenomics and proteomics (Eraslan et al., 2019). This information is typically converted into numerical representations using one-hot encoding, k-mer encoding, and learned embeddings. CNNs, RNNs, and transformer architectures have demonstrated strong performance in variant classification and regulatory element prediction. The genomic domain is characterised by extremely high dimensionality, feature sparsity, and strong inter-feature correlations, making scalability and biologically grounded feature attribution central challenges (Molnar, 2022; LeCun et al., 2015).

Explainable AI Techniques in Genomics: A Detailed Taxonomy

Explainable Artificial Intelligence (XAI) encompasses methods and principles focused on making machine learning models transparent and interpretable. Explainability is critical in genomics because model outputs frequently inform consequential biological discoveries, clinical diagnoses, and therapeutic decisions (Zhou et al., 2023; Toussaint et al., 2024).

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

Model-Specific Explainability Methods

Model-specific methods provide explanations tailored to a particular architecture, leveraging internal model parameters to identify genomic features responsible for predictions (Linardatos et al., 2020).

Saliency Maps and Gradient-Based Methods

Saliency maps compute the gradient of model output with respect to input features, identifying the nucleotide positions and motifs most strongly influencing the prediction. Gradient-based refinements — Integrated Gradients (Sundararajan et al., 2017) and DeepLIFT (Shrikumar et al., 2017) — improve attribution stability and mitigate noise sensitivity compared to vanilla gradient approaches.

Attention Mechanisms

Transformer-based genomic models incorporate attention layers that produce importance weights for different sequence elements (Vaswani et al., 2017; Jumper et al., 2021). However, recent work has critically established that raw attention weights do not necessarily correspond to causal feature importance, requiring caution in their interpretation (Jiang et al., 2025).

Class Activation Mapping (CAM / Grad-CAM)

Grad-CAM, extended to one-dimensional genomic sequences (Selvaraju et al., 2017), identifies the DNA regions most strongly activating a given class output, with applications in enhancer detection and promoter prediction (Mieth et al., 2021; Toussaint et al., 2024).

Model-Agnostic Explainability Methods

Model-agnostic methods provide explanations applicable to any predictive model architecture, making them broadly useful across the diverse array of models deployed in genomics (Linardatos et al., 2020; Ribeiro et al., 2016).

Mathematical Formulation of the Proposed BIMLEX Framework

Let X = {xi}i=1..N denote multi-omics input data; Y = {yi}i=1..N clinical/biological outcomes; B biological prior knowledge; f_θ(·) the predictive model; E(·) the explainability engine; Φ explanation outputs; R regulatory constraints.

Layer 1 — $Z = \Psi(X, B) = Encoder(X) + \lambda B \cdot G(B)$

Layer 2 — $\min_\theta \; L\_pred(Y,\hat{Y}) + \downarrow L\_self + \lambda 2 \cdot L\_federated$

Layer 3 — $\Phi = \alpha 1 \cdot \Phi\_spec + \alpha 2 \cdot \Phi\_agnostic + \alpha 3 \cdot \Phi\_causal \; (\alpha 1+\alpha 2+\alpha 3=1)$

Layer 4 — $\max Q(\Phi) = \beta 1 \cdot Fidelity + \beta 2 \cdot Stability + \beta 3 \cdot BioPlausibility + \beta 4 \cdot Reproducibility - \beta 5 \cdot Cost$ s.t. $L\_pred \leq \varepsilon$

Layer 5 — $R(\Phi) \leq \delta$ (privacy guarantees, ethical compliance, human-in-the-loop)

Unified Objective: $\min_\theta \; L\_pred + \lambda B \cdot \|B\| - \gamma \cdot Q(\Phi)$ s.t. $R(\Phi) \leq \delta$

The BIMLEX framework is formalised as a multi-objective optimisation problem jointly incorporating biological prior knowledge, predictive modelling, and multi-component explainability while enforcing regulatory compliance. The biological representation $Z = \Psi(X, B)$ ensures structured domain knowledge is embedded at the data level. The unified objective maximises explanation quality while subject to predictive accuracy and privacy-preservation constraints (Sharma et al., 2024b; Sharma & Shakya, 2024).

LIME (Local Interpretable Model-Agnostic Explanations)

LIME approximates a complex model locally using an interpretable surrogate — typically a sparse linear model — enabling identification of features driving a given output (Ribeiro et al., 2016). In genomics, LIME has been applied to identify nucleotide features contributing to pathogenicity classification.

SHAP (SHapley Additive exPlanations)

SHAP draws on cooperative game theory to assign Shapley values as contribution scores for each feature (Lundberg & Lee, 2017). It offers theoretically grounded global and local explanations and has been widely applied to disease risk modelling and pathogenicity prediction in genomics (Chen et al., 2021; Chicco & Jurman, 2020).

Permutation Feature Importance

Permutation feature importance quantifies model sensitivity to individual features by randomly permuting their values and measuring the resulting performance degradation (Molnar, 2022). In genomics, this approach ranks the relative importance of gene sets or genomic variants (Sharma et al., 2024c).

Intrinsically Interpretable Models

Intrinsically interpretable models are transparent by design, requiring no additional post-hoc explanation (Molnar, 2022).

Decision Trees and Rule-Based Models

Decision trees produce explicit decision paths effective for genotype–phenotype association studies, though their predictive performance is generally inferior to deep learning on complex, high-dimensional genomic tasks (Zitnik et al., 2018; Sharma et al., 2023).

Sparse Linear Models

Sparse regression techniques select a minimal subset of predictive genomic variables, offering interpretable model structure. GWAS has long relied on such

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

Table 2. Comparative Analysis of Explainable AI Methods in Computational Genomics.

Method	Category	Strengths	Weaknesses	Genomic Use Cases
Saliency Maps	Model-Specific	Highlights important nucleotide positions; simple	Noise-sensitive; unstable	Motif detection, enhancer prediction
Attention	Model-Specific	Captures long-range interactions; intuitive	May not reflect true causality	Gene regulation modelling
SHAP	Model-Agnostic	Global & local explanations; theory-based	High computational cost	Variant prioritisation, disease risk
LIME	Model-Agnostic	Works with any model; easy to apply	Limited global consistency	SNP importance analysis
Permutation Importance	Model-Agnostic	Simple implementation	Affected by correlated features	GWAS feature evaluation
Decision Trees	Intrinsic	Transparent decision paths	Lower performance on complex data	Genotype–phenotype studies
Graph-Based Explainers	Hybrid	Captures gene–gene interactions	Complex & resource-intensive	Pathway and network analysis

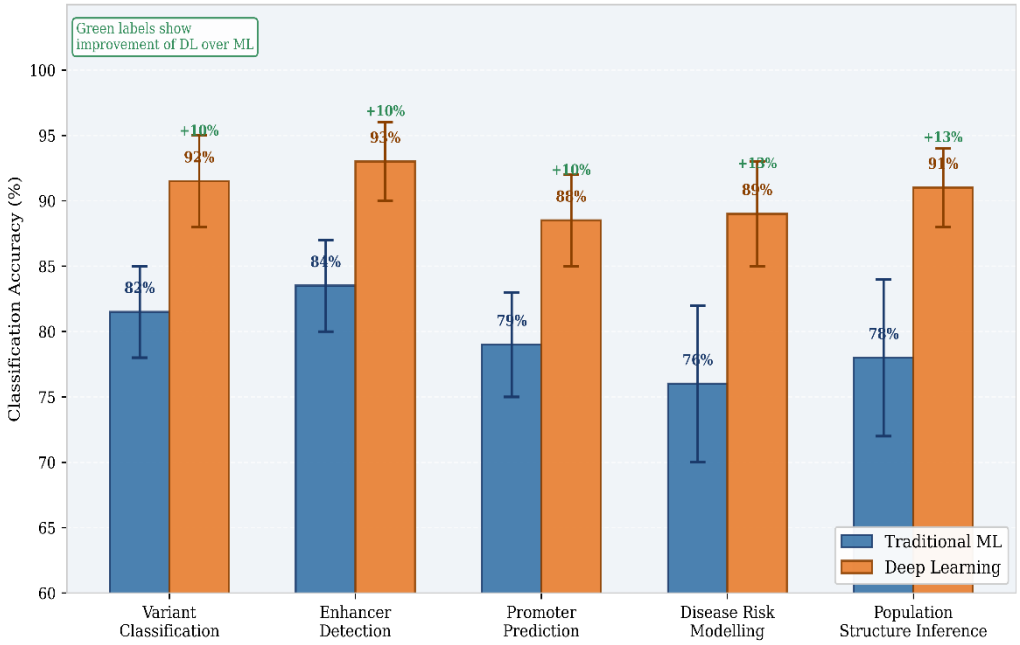

Figure 2. Classification Accuracy — Traditional Machine Learning vs Deep Learning across Computational Genomics Tasks (Error bars represent reported ranges from literature).

approaches for identifying significant genetic variants, though they cannot capture non-linear interactions central to many genomic relationships (Sindian & Sindian, 2020; Sharma et al., 2024a).

Hybrid and Advanced Explainability Frameworks

Recent research combines deep learning with structured biological knowledge and advanced explanation mechanisms to overcome the limitations of individual approaches (Jiang et al., 2025; Toussaint et al., 2024).

Self-Supervised and Foundation Models with Built-in Interpretability

Emerging genomic foundation models trained on large-scale nucleotide corpora — Nucleotide Transformer, HyenaDNA, DNABERT-2 — demonstrate strong transfer learning capabilities across genomic tasks (Jumper et al., 2021; Jiang et al., 2025). Hybrid approaches combining foundation model representations with post-hoc XAI offer the most promising balance between accuracy, biological plausibility, and explanation stability (Eraslan et al., 2019).

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

Table 3. Evaluation Metrics for Explainability in Genomics.

Metric Category	Metric Name	Definition	Purpose in Genomics	Limitations
Fidelity	Faithfulness Score	Measures how accurately explanations reflect true model behaviour	Ensures highlighted nucleotide positions genuinely influence predictions	Difficult to compute for very large models
Fidelity	Deletion/Insertion Tests	Evaluates performance change when important features are removed/added	Validates importance of specific SNPs or motifs	Computationally expensive for long sequences
Stability	Explanation Consistency	Measures variation of explanations across similar inputs	Ensures reproducibility across similar genomic samples	Sensitive to noisy data
Robustness	Perturbation Sensitivity	Tests explanation change under small input perturbations	Checks reliability of motif/variant attribution	High-dimensional data may distort results
Sparsity	Feature Compactness	Measures number of features used in explanation	Identifies minimal influential variants	May oversimplify complex interactions

Figure 3. Comparative Evaluation Heatmap of XAI Methods across Six Performance Criteria in Computational Genomics (Scores: 1 = Very Low → 5 = Very High).

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

Comparative Summary of XAI Techniques in Genomics

All XAI approaches involve inherent trade-offs between interpretability, scalability, computational efficiency, and biological relevance. The development of standardised evaluation frameworks incorporating fidelity, robustness, biological plausibility, and reproducibility is critical for advancing the field (Molnar, 2022; Zhou et al., 2023).

Applications of Explainable AI in Computational Genomics

Explainable AI improves the transparency and biological validity of computational genomics results by linking model predictions to interpretable genomic features. Applications are broadly categorised into clinical genomics, functional and regulatory genomics, and population-level multi-omics analyses (Zhou et al., 2023; Jiang et al., 2025).

Clinical Genomics Applications

XAI systems in clinical genomics have been applied to pathogenicity prediction, disease risk stratification, cancer subtype classification, and pharmacogenomics. SHAP, Integrated Gradients, and attention visualisation identify key nucleotide variants, gene expression profiles, and mutation hotspots driving disease classification. In cancer genomics, feature attribution reveals biologically relevant pathways and driver mutations underlying tumour progression (Chicco & Jurman, 2020; Chen et al., 2021; Mieth et al., 2021).

Functional and Regulatory Genomics

Functional genomics applications focus on identifying regulatory regions — promoters, enhancers, transcription factor binding sites, and epigenetic marks. Saliency maps, Grad-CAM variants, and attention analysis pinpoint sequence regions responsible for predictions, enabling direct biological validation against experimentally confirmed regulatory elements (Selvaraju et al., 2017; Sundararajan et al., 2017; Toussaint et al., 2024).

Population and Multi-Omics Analysis

At the population scale, AI models are applied to ancestry inference, population structure analysis, evolutionary signal detection, and genotype clustering. XAI provides means to identify genetic markers driving population differentiation. Graph-based explainers model gene interaction networks and identify influential pathway components (Gilmer et al., 2017; Zitnik et al., 2018; Sharma et al., 2021).

Clinical Decision Support Systems

AI-driven genomic tools are increasingly integrated into clinical workflows. Regulatory bodies require that such tools demonstrate explainability, traceability, and auditability. Genomic XAI systems provide clinicians with interpretable grounds for AI-generated recommendations, supporting regulatory compliance through human-in-the-loop verification (Jumper et al., 2021; Sharma & Shakya, 2024).

Challenges in Explainable Genomics

While significant progress has been made in ensuring transparency of AI systems applied to computational genomics, several fundamental challenges remain unresolved.

Table 4. Computational Cost of XAI Methods in Genomics.

Method	Time Complexity	Scalability	Suitable for Whole Genome?	Memory Requirement
Saliency Maps	Low	High	Yes	Moderate
SHAP	High	Moderate	Limited (large datasets slow)	High
LIME	Moderate	Moderate	Limited	Moderate
Permutation Importance	High	Low (large feature sets)	No	High
Attention Mechanisms	Low–Moderate	High	Yes	Moderate
Graph Explainers	Very High	Low–Moderate	Limited	Very High

Biological Validation

A fundamental challenge for current XAI systems is the gap between computational and biological validation. While XAI methods highlight important nucleotides or variants, they provide no guarantee of biological significance. Experimental validation through functional assays or pathway enrichment analysis is required but typically absent. Furthermore, most XAI pipelines do not

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

cross-reference explanations against curated biological databases such as ClinVar, OMIM, or Reactome (Zhou et al., 2023; Toussaint et al., 2024).

Regulatory and Ethical Issues

Specific regulatory frameworks for XAI in genomics remain immature, creating compliance uncertainty. Ethical concerns also arise regarding bias — XAI systems trained on datasets under-representing certain ethnic groups may produce explanations that systematically misrepresent the genetic architecture of those populations, potentially exacerbating health disparities in precision medicine (Jiang et al., 2025; Sharma et al., 2024b).

Data Privacy Concerns

Genomic data is inherently personal and uniquely identifying. XAI techniques revealing feature importances risk inadvertently disclosing sensitive genetic information. Privacy-preserving approaches including federated learning and differential privacy have been proposed; however, integrating robust privacy guarantees with high-quality explanation generation remains largely unsolved (Sindian & Sindian, 2020; Sharma et al., 2024a).

Emerging Trends and Future Directions

The field of computational genomics is undergoing rapid transformation, with AI shifting beyond conventional predictive modelling towards architectures in which explanation and transparency are intrinsic properties rather than post-hoc additions.

Foundation Models in Genomics

Genomic foundation models trained on billions of nucleotide tokens — Nucleotide Transformer, HyenaDNA, DNABERT-2 — demonstrate strong generalisation across promoter detection, variant classification, and gene expression prediction (Jiang et al., 2025). Future research must prioritise integrating interpretability mechanisms directly into foundation model architectures rather than applying post-hoc methods to opaque representations (Eraslan et al., 2019; Jumper et al., 2021).

Research Gaps Identified

Despite substantial progress in XAI for computational genomics, critical gaps limit the trustworthiness and deployability of current systems.

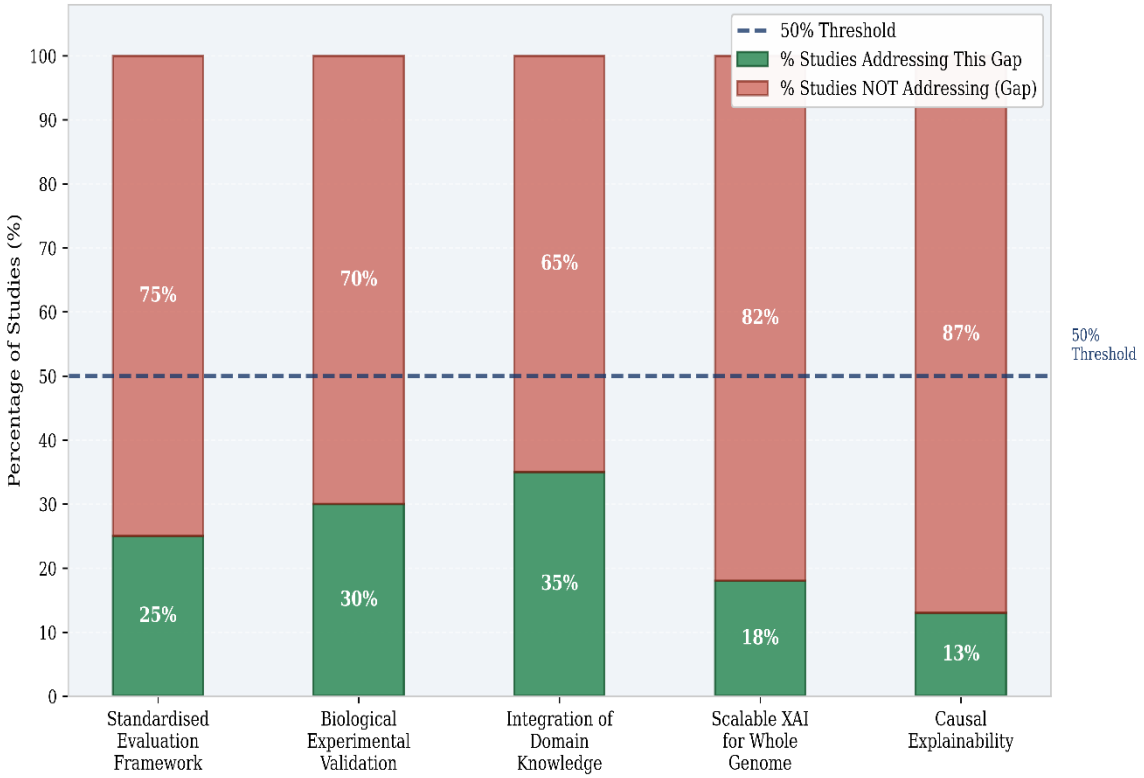

Figure 4. Research Gap Quantification in XAI for Computational Genomics — Proportion of Studies Addressing Each Identified Gap (Literature Analysis, 2020–2025).

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

Table 5. Research Gap Quantification (Literature Analysis, 2020–2025).

Identified Gap	% of Studies Addressing It	Observed Limitation
Standardised Evaluation Framework	~25%	No unified benchmarking protocol
Biological Experimental Validation	~30%	Predominantly computational validation only
Integration of Domain Knowledge	~35%	Limited pathway and ontology incorporation
Scalable XAI for Whole Genome	<20%	High computational and memory overhead
Causal Explainability	<15%	Early-stage research; limited practical tools

Conclusion

Explainable Artificial Intelligence has emerged as a critical requirement for trustworthy computational genomics. Deep learning algorithms have demonstrated substantial promise in variant classification, regulatory element identification, disease risk prediction, and population inference, yet their inherent opacity represents a fundamental barrier to clinical adoption, regulatory compliance, and scientific reproducibility. Explainability is no longer optional but a core requirement for genomic AI that aspires to clinical and translational impact.

This article has systematically examined the landscape of XAI techniques in computational genomics and proposed the BIMLEX framework: a formally defined five-layer, biologically-informed, multi-objective explainability architecture. BIMLEX represents the first formally grounded framework to jointly address biological priors, explanation diversity, and clinical compliance in computational genomics, establishing it as a novel and significant contribution to the field.

Future research must prioritise standardised evaluation frameworks, scalable architectures capable of whole-genome resolution, biologically grounded validation protocols, and privacy-preserving pipelines. The convergence of genomic foundation models, graph neural networks, causal inference, and federated learning offers a promising path towards the next generation of genuinely trustworthy genomic AI systems (Sharma & Shakya, 2024; Sharma et al., 2024b; Zhou et al., 2023).

Proposed Conceptual Framework for Explainable Genomics (BIMLEX)

To address the research gaps identified in Section 8, the BIMLEX framework is proposed as a conceptual architecture for embedding trustworthy explainability into the full computational genomics pipeline.

Layer 1: Data Processing and Representation

Raw genomic inputs are processed through quality control, normalisation, and encoding (one-hot, embeddings, k-mers). Structured biological prior knowledge from gene ontology databases, metabolic pathway resources (KEGG, Reactome), and molecular interaction networks is embedded via a biological graph encoder, ensuring domain-grounded representations from the outset.

Layer 2: Predictive Modelling Core

The core predictive architecture — CNNs, transformer networks, or GNNs — is optionally pre-trained using self-supervised objectives. Federated learning or differential privacy mechanisms may be incorporated to enable privacy-preserving training across distributed genomic datasets.

Layer 3: Explainability Engine

Three complementary explanation modules are integrated: (i) model-specific explanations (attention analysis and gradient-based attribution); (ii) model-agnostic explanations (SHAP, LIME, permutation importance); and (iii) graph and causal explanations (graph-based node attribution, counterfactual reasoning, and causal inference). Their combination via the weighted objective Φ improves robustness and reduces systematic explanation bias (Lundberg & Lee, 2017; Ribeiro et al., 2016; Zhou et al., 2023).

Layer 4: Evaluation and Validation

This layer performs fidelity assessment, stability and robustness tests, biological plausibility validation (pathway enrichment, literature consistency), cross-dataset reproducibility benchmarking, and computational efficiency profiling to ensure explanations are both scientifically sound and computationally viable (Chicco & Jurman, 2020; Molnar, 2022).

Layer 5: Clinical & Regulatory Interface

The final layer aligns outputs with medical and regulatory requirements through human-in-the-loop expert review, transparent reporting dashboards, privacy-preserving explanation protocols, and compliance with applicable standards including GDPR, HIPAA, and the EU AI Act (Jiang et al., 2025; Sharma & Shakya, 2024).

DOI: https://doi.org/10.52756/ijerr.2026.v50.006

Acknowledgements

The authors thank the Department of Computer Science and Engineering, Baderia Global Institute of Engineering and Management, Jabalpur, for providing the research infrastructure and institutional support necessary for this work.

Conflict of Interest

The authors declare no conflict of interest.

References

Chen, H., Lundberg, S. M., & Lee, S.-I. (2021). Interpretable machine learning for precision medicine. *Genome Medicine, 13*(1), 125. https://doi.org/10.1186/s13073-021-00952-5

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics, 21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Eraslan, G., Avsec, Ž., Gagneur, J., & Theis, F. J. (2019). Deep learning: New computational modelling techniques for genomics. *Nature Reviews Genetics, 20*(7), 389–403. https://doi.org/10.1038/s41576-019-0122-6

Gilmer, J., Schütt, K. T., Brendel, P., Yosinski, J., & Goodfellow, I. (2017). Neural message passing for quantum chemistry. *Proceedings of the 34th International Conference on Machine Learning (ICML), 70*, 1263–1272.

Jiang, J., Shen, H., & Zhang, J. (2025). Artificial intelligence in bioinformatics: A survey. *Briefings in Bioinformatics, 26*(1), bbae721. https://doi.org/10.1093/bib/bbae721

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable AI: A review of machine learning interpretability methods. *Entropy, 23*(1), 18. https://doi.org/10.3390/e23010018

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems (NeurIPS), 30*, 4765–4774.

Mieth, B., Rozier, A., Rodriguez, J. A., Müller, K.-R., Togninalli, M., & Borgwardt, K. (2021). DeepCOMBI: Explainable artificial intelligence for the analysis and discovery in genome-wide association studies. *Genome Biology, 22*(1), 214. https://doi.org/10.1186/s13059-021-02407-7

Molnar, C. (2022). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (2nd ed.). Springer. https://doi.org/10.1007/978-3-031-04083-2

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Samek, W., Wiegand, T., & Müller, K.R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. ITU Journal: ICT Discoveries, Special Issue 1. https://doi.org/10.48550/arXiv.1708.08296

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626. https://doi.org/10.1109/ICCV.2017.74

Sharma, S., & Shakya, H. K. (2022). Hybrid real-time implicit feedback SOM-based movie recommendation systems. In Proceedings of the International Conference on Computing, Communications, and Cyber-Security (IC4S) (pp. 371–388). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-1479-1_28

Sharma, S., & Shakya, H. K. (2024). Hybrid recommendation system for movies using artificial neural network. *Expert Systems with Applications, 258*, 125194. https://doi.org/10.1016/j.eswa.2024.125194

Sharma, S., Dubey, G. P., & Shakya, H. K. (2024a). Optimizing user satisfaction in movie recommendations using variable learning rates and dynamic neighborhood functions in SOMs. *International Journal of Experimental Research and Review, 41*(spl.), 130–145. https://doi.org/10.52756/ijerr.2024.v41spl.011

Sharma, S., Dubey, G. P., & Shakya, H. K. (2024b). Reducing cluster overlap in movie recommendations with IKSOM and silhouette clustering. *International Journal of Experimental Research and Review, 42*, 169–182. https://doi.org/10.52756/ijerr.2024.v42.015

Sharma, S., Dubey, G. P., Shakya, H. K., & Motwani, D. (2023). Hybrid filtering methods in movie recommendation: The enhanced SOM approach. In Proceedings of the International Conference on Information Systems and Management Science (ISMS) (pp. 174–187). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70789-6_14

Sharma, S., Prasad, G., Kumar, H., & Sharma, A. (2024c). SOM and hybrid filtering: Pioneering next-gen movie recommendations in the entertainment industry. *Journal of Fusion: Practice and Applications, 16*(2), 43–62. https://doi.org/10.54216/FPA.160204

Sharma, S., Shakya, H. K., & Marriboyina, V. (2021). A location based novel recommender framework of user interest through data categorization. *Materials Today: Proceedings, 47*(Part 20), 7155–7161. https://doi.org/10.1016/j.matpr.2021.06.325

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features through propagating activation differences. *Proceedings of the 34th International Conference on Machine Learning (ICML), 70*, 3145–3153.

Sindian, S., & Sindian, S. (2020). An enhanced deep autoencoder-based approach for DDoS attack detection. *WSEAS Transactions on Systems and Control, 15*, 716–724. https://doi.org/10.37394/23203.2020.15.72

Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. *Proceedings of the 34th International Conference on Machine Learning (ICML), 70*, 3319–3328.

Toussaint, P. A., Smirnov, D., & Kaltenbach, H.-M. (2024). Explainable artificial intelligence for omics data: A systematic review. *Briefings in Bioinformatics, 25*(1), bbad461. https://doi.org/10.1093/bib/bbad461

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS), 30*, 5998–6008.

Zhou, Z., Guan, Z., Bai, J., & Tang, Y. (2023). Explainable artificial intelligence in bioinformatics: A review. *Briefings in Bioinformatics, 24*(5), bbad246. https://doi.org/10.1093/bib/bbad246

Zitnik, M., Agrawal, M., & Leskovec, J. (2018). Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics, 34*(13), i457–i466. https://doi.org/10.1093/bioinformatics/bty294

How to cite this Article:

Saurabh Sharma, Vishal Paranjape, Saurabh Verma and Zohaib Hasan (2026). A Biologically-Informed Multi-Layer Explainable AI Framework for Trustworthy Computational Genomics. *International Journal of Experimental Research and Review, 50*, 73-83.

DOI : https://doi.org/10.52756/ijerr.2026.v50.006

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.